

Rec'd EPO 13 APR 2005

10/531197



Europäisches
Patentamt

European
Patent Office

Office européen
des brevets

IP & S-DE zugestellt
am 17. Okt. 2003
Frist

Bescheinigung

Certificate

Attestation

REC'D 31 OCT 2003

WIPO PCT

Die angehefteten Unterla-
gen stimmen mit der
ursprünglich eingereichten
Fassung der auf dem näch-
sten Blatt bezeichneten
europäischen Patentanmel-
dung überein.

The attached documents
are exact copies of the
European patent application
described on the following
page, as originally filed.

Les documents fixés à
cette attestation sont
conformes à la version
initialement déposée de
la demande de brevet
européen spécifiée à la
page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

02102454.2

1B/03/4429

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk



Anmeldung Nr:
Application no.: 02102454.2
Demande no:

Anmeldetag:
Date of filing: 16.10.02
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

Koninklijke Philips Electronics N.V.
Groenewoudseweg 1
5621 BA Eindhoven
PAYS-BAS

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se référer à la description.)

Directory assistant method and apparatus

In Anspruch genommene Priorität(en) / Priority(ies) claimed / Priorité(s)
revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

G06F12/00

Am Anmeldetag benannte Vertragsstaaten/Contracting states designated at date of
filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LI LU MC NL PT SE SK TR

DESCRIPTION

DIRECTORY ASSISTANT METHOD AND APPARATUS

5

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention relates to a directory assistant method and apparatus, and
 10 more particularly, to a directory assistant method and apparatus in an automatic
 dialogue telecommunications system.

Description of the Related Art

The directory assistant (DA) system, providing telephone numbers to customers,
 15 is an important telecommunications business. For example, the Kelsey Group
 estimates that telecom companies worldwide collectively receive more than 516
 millions DA calls per month, almost all of which are currently handled by
 operators. Automating this service using speech recognition is a large market
 opportunity.

20

The conventional DA system is implemented by using a restricted dialogue.
 Traditionally, it first asks a user to say the name of the person to be reached, then
 uses a speech recognizer to locate several candidates from a directory database. If
 the candidates are too many, the DA system further asks the user to spell the
 25 name of the desired person or to provide extra information, for example, the name
 of the street where the desired person lives. In this way, the rang of the
 candidates can be further narrowed down. Finally, the DA system asks the user to
 choose the right one by answering a corresponding number or just "yes/no". This
 DA system works well for a small Western DA system. But it may not work well
 30 for a large-scale directory assistant system having, for example, 12,000,000
 entries used in a large city, since the above-mentioned input information is not
 sufficient to differentiate all possible candidates.

The same system does not work well for a large-scale Chinese DA system, either. The input information is not sufficient to differentiate all possible candidates due to the following specific features. First of all, Chinese is a monosyllabic language. Each word of Chinese contains exactly one syllable. There are more than 13000 commonly used words and only 1300 legal syllables. On average, there are about 10 homophones for each syllable. Secondly, a Chinese name is usually shorter than a Western name. The Chinese name usually has three syllables only. Moreover, there are about two hundred family names (surnames) frequently used by billions of Chinese. Therefore, more information is needed to solve the ambiguities in a Chinese DA system. Thirdly, Chinese is an ideographic language. Chinese usually introduce their names to other people by describing their name word by word and by some commonly used phrases. There is no easy and standard way to "spell" or "compose" Chinese words. Therefore, the performance of current DA systems is not satisfactory, especially the Chinese DA systems.

SUMMARY OF THE INVENTION

It is an object of the present invention to provide a directory assistant method and apparatus for providing desired directory entry information. The directory assistant method and apparatus use natural language dialogue system to ask users to describe the desired directory entry and then use the relevance knowledge databases to parse and understand these descriptions and interpret their meanings. Finally, the directory assistant method and apparatus integrate all available information from several dialogues turns, directory database and relevant knowledge database, and then provide the users' desired directory entry information.

It is another object of the present invention to provide a computer program product residing on a computer readable medium having a plurality of instructions stored thereon which, when executed by a processor, cause the

processor to provide desired directory entry information.

BRIEF DESCRIPTION OF THE DRAWINGS

5 The invention will become more fully understandable from the detailed description given below and the accompanying drawings, which are given by way of illustration only and thus are not imitative, wherein:

Fig. 1 illustrates a block diagram of the directory assistant apparatus of the present invention;

10 Fig. 2 illustrates the generation of name description grammar rules from frequently used templates and frequently used words of the present invention; and

15 Fig. 3 illustrates a flow chart of the directory assistant method of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

Referring to Fig. 1, the directory assistant apparatus of the present invention, such as a Mandarin directory assistant apparatus, comprises a database 30 for storing
20 directory entry information, grammar rules and concept sequences; an acoustic recognition unit 10 for receiving a speech signal describing the desired directory entry, recognizing the speech signal and generating recognized word sequences; a speech interpreting unit 20 for interpreting the recognized word sequences by
25 using a predetermined grammar rule and relevant information thereof stored in the database 30 to form concept sequences and interpreting the concept sequences according to semantic meaning and relevant information thereof stored in the database and current system status thereof, thereby generating at least one candidate by using one of maximum a posterior probabilities and maximum likelihood criterion for the desired directory entry, in addition, the speech
30 interpreting unit 20 further updates the system status; a look up unit 40 for looking up at least one directory entry information corresponding to the at least

one candidate from the database 30; and an output unit 60 (such as a speech output unit) for outputting the at least one directory entry information located.

5 The directory assistant apparatus of the present invention further comprises a question generator 60 for generating a question to request more information, wherein the question is one of requesting a user to supply more information, listing-based confirmation and open-question confirmation. The listing-based confirmation is used when the potential candidates are in a limit of numbers, or the probability of the top one is far from those of the others. The open-question
10 confirmation is used when the most popular description way in name database to ask users for confirmation, for example: you mean 李登輝的李 (the same Li3 as in Li3 Deng1 Hue1).

The acoustic recognition unit 10 further comprises a speech recognizer 11 for
15 recognizing the input speech signal and generating recognized word sequences; a confusion analyzer 12 for expanding the recognized word sequences according to a confusion table 13, wherein the confusion table 13 is pre-trained and comprises all confusable words, their corresponding correct ones and occurring probabilities; and a confidence measurement unit 14 for filtering out confusable word pairs
20 according to a confidence table 15.

The database 30 comprises a relevant knowledge database 31 and a directory database 32. The relevant knowledge database 31 comprises words and using frequencies thereof, ways to describe the words, grammar rules, attributes and
25 corresponding using frequencies, communication concepts and their frequencies of usage, corresponding grammar rules, semantic meaning and frequencies of usage, while the directory database 32 comprises a plurality of entries, wherein each entry comprises name, telephone number, relevant information and frequencies of usage.

30

In the relevant knowledge database 31, popular words in names are stored with several popular descriptive ways, wherein the grammar rule and concept

sequences are used to describe the desired directory entry comprising entry name or at least one word of entry name and relevant information thereof, and the grammar rule is generated by frequently used grammar templates and frequently used words. The grammar templates are generated by one of frequently used
 5 nouns, names of famous people, idioms, character strokes, letters, words, and character roots, etc. For examples, words in names can be described as follows:

- famous family name description, like 李登輝的李 (same Li3 as in Li3 Deng1 Hue1).
- 10 - famous name description, like 李登輝的李 (same Li3 as in Li3 Deng1 Hue1).
- Common used word, phrase and especially four-word Chinese idioms, like 趙錢孫李的趙 (same Zhau4 as in Zhau4 Chien2 Sun1 Li3).
- common used writing/strokes description, like 三橫一豎王 (Wang2 that has
 15 three horizontal lines and one vertical line); or 耳東陳 (Chen2 that has an ear and a east).

Fig. 2 illustrates the generation of name description grammar rules from frequently used templates and frequently used words of the present invention.
 20 The present invention first builds a database which collects the name description grammar rules and their corresponding semantic tags.

There are two ways to build the database. The first way is to collect as many names as possible and their corresponding character descriptions. From this
 25 database, we have found name description grammar rules and their probability statistics such as LN (descriptions of the last name) 84, FN1 (descriptions of the first word of the first name) 85 and FN2 (description of the second word of the first name) 86.

30 The second way is to find frequently used grammar templates from a small database of name descriptions (for example the database mentioned above). Then

we use the found grammar templates and frequently used words to generate the necessary grammar rules. For example, we have found that the most popular ways to describe the words of names are:

- 5 - Frequently used nouns (FNoun) 81;
- Names of famous people (FName) 82;
- Idioms (CI);
- Character Strokes (CS);
- Frequently used foreign words (FW);
- 10 - Character roots (CR) 83;
-
- Other irregular way (OW).

We can then build the necessary grammar rules by combining these grammar
15 templates and frequently used words (collected from dictionary, internet,
 newspaper, etc.).

Referring to Fig. 3, the directory assistant method of the present invention, such
as a Mandarin directory assistant method, is described as follows:

20 The method first prompts a question to ask the user to speech input the desired
 entry (100); then receives a speech signal describing the desired directory entry
 (110); recognizes the speech signal and generating recognized word sequences,
 expands the recognized word sequences according to a confusion table and filters
25 out confusable word pairs according to a confidence table (120); interprets the
 recognized word sequences by using a predetermined grammar rule and relevant
 information thereof stored in a database to form concept sequences (130);
 interprets the concept sequences according to semantic meaning and relevant
 information thereof stored in the database and current system status thereof,
30 thereby generating at least one candidate by using one of maximum a posterior
 probabilities and maximum likelihood criterion for the desired directory entry and

updates the system status; looks up at least one directory entry information corresponding to the at least one candidate from the database and generates a question to request more information if there are uncertainties (150); outputs the at least one directory entry information located (160); and confirms the located
5 directory entry information and repeats the above steps until the desired directory entry information is located (170).

The above-mentioned method can be implemented by computer program instructions. The computer program instructions can be loaded into a computer
10 or other programmable processing devices to perform the functions of the method illustrated in Fig. 3. The computer program instructions can be stored in a computer program product or computer readable medium. Examples of a computer program product or computer readable medium includes recordable-type medium such as a magnetic tape, a floppy disc, a hard disc drive, a RAM, a
15 ROM and an optical disc and transmission-type medium such as digital and analog communication links.

The present invention is directed to understanding ways of describing words in names, building up a relevant knowledge database to store ways of describing
20 words in names and using the database as grammar rules to parse input speech. By this new architecture, the present invention can use the natural language dialogues system to ask the user to describe the words of names when there are still uncertainties. The present invention then uses the relevance knowledge database to parse and understand these descriptions and interpret their meaning.
25 Finally, the present invention combines all available information to narrow down the range of possible candidates and finally locates the correct directory entry. Although part of the present invention is described by using Chinese words as an example, the present invention can also be applied to other languages. For example, famous family name description, like 李登輝的李 (same Li3 as in Li3
30 Deng1 Hue1), can be changed to "Bush as in George Bush."

Although the present invention and its advantages have been described in detail, it should be understood that various changes, substitutions and alternations can be made herein without departing from the spirit and scope of the invention as defined by the appended claims.

PRIOR ART REFERENCES

1. Nick Wang and Leo Liao, Chinese name recognition, *invention disclosure*, submitted to Philips Corporate Intellectual Property, 11 Oct. 2000
52. Andreas Kellner, Bernhard Rueber, Frank Seide, and Bach-Hiep Tran., "PADIS -- an Automatic Telephone Switchboard and Directory Information System", *Speech Communication*, 23:95--111, Oct 1997
3. Bernd Souvignier, Andreas Kellner, Bernhard Rueber, Hauke Schramm, and Frank Seide., "The Thoughtful Elephant: Strategies for Spoken Dialog Systems". *IEEE Transactions on Speech and Audio Processing*, 8(1):51--62, January 2000
- 10 4. Georg Rose, "PADIS-XXL, Large Scale Directory Assistant," PowerPoint Slide, Man-Machine Interface, 2000.
5. Andreas Meylahn, "SpeechFinder for SpeechPerl 2000 - Developer's Guide v2.0", Philips Speech Processing - Aachen, Germany, 21 Sep. 2000
156. Yen-Ju Yang (楊燕珠), "Statistics-based spoken dialogue modeling and its applications," PhD's Thesis, National Taiwan University, Taiwan, 1999.
7. Jan Kneissler and Dietrich Klakow, "Speech Recognition for Huge Vocabularies by Using Optimized Sub-Word Units"
8. Dietrich Klakow, Georg Rose and Xavier Aubert, "OOV-Detection in Large Vocabulary System Using Automatically Defined Word-Fragments as Fillers", *Proc. EUROSPEECH*, Vol. 1, 1999, pp. 49-53.
- 20 9. Alex Weibel, Petra Geutner, Laura Mayfield Tomokiyo, Tanja Schultz and Monika Woszcyna, "Multilinguality in Speech and Spoken Language Systems," *Proceedings of the IEEE*, Vol. 88, No. 8, August 2000, pp. 1297-1313.

CLAIMS

1. A directory assistant method for providing desired directory entry information, comprising the steps of:

- 5 (a) receiving a speech signal describing the desired directory entry;
- (b) recognizing the speech signal and generating recognized word sequences;
- (c) interpreting the recognized word sequences by using a predetermined grammar rule and relevant information thereof stored in a database to form concept sequences;
- 10 (d) interpreting the concept sequences according to semantic meaning and relevant information thereof stored in the database and current system status thereof, thereby generating at least one candidate for the desired directory entry;
- 15 (e) looking up at least one directory entry information corresponding to the at least one candidate from the database; and
- (f) outputting the at least one directory entry information located.

20 2. The method as claimed in Claim 1, further comprising the steps of user's correction or confirmation and repeating the steps (a) to (f) until the desired directory entry information is located.

3. The method as claimed in Claim 1, wherein the step (a) further comprises the step of system prompting before receiving a speech signal.

25

4. The method as claimed in Claim 1, wherein the predetermined grammar rule and concept sequences are used to describe the desired directory entry comprising entry name or at least one word of entry name and relevant information thereof.

30

5. The method as claimed in Claim 1, wherein the predetermined grammar rule is generated by frequently used grammar templates and frequently used words.

5 6. The method as claimed in Claim 5, wherein the grammar templates are generated by one of frequently used nouns, names of famous people, idioms, character strokes, letters, words, and character roots.

10 7. The method as claimed in Claim 1, wherein the database comprises a relevant knowledge database and a directory database.

8. The method as claimed in Claim 1, wherein the step (b) further comprises the step of expanding the recognized word sequences according to a confusion table.

15 9. The method as claimed in Claim 8, wherein the confusion table is pre-trained and comprises all confusable words, their corresponding correct ones and occurring probabilities.

20 10. The method as claimed in Claim 1, wherein the step (b) further comprises the step of confidence measurement for filtering out confusable word pairs according to a confidence table.

25 11. The method as claimed in Claim 1, wherein the step (d) further comprises the step of updating the system status.

12. The method as claimed in Claim 1, wherein the step (e) further comprises the step of generating a question to request more information.

30 13. The method as claimed in Claim 12, wherein the question is one of requesting a user to supply more information, listing-based confirmation and open-question confirmation.

14. The method as claimed in Claim 7, wherein the relevant knowledge database comprises words and using frequencies thereof, ways to describe the words, grammar rules, attributes and corresponding frequencies of usage.

5 15. The method as claimed in Claim 7, wherein the relevant knowledge database comprises communication concepts and frequencies of usage thereof, corresponding grammar rules, semantic meaning and frequencies of usage.

10 16. The method as claimed in Claim 7, wherein the directory database comprises a plurality of entries, wherein each entry comprises name, telephone number, relevant information and using frequency.

15 17. The method as claimed in Claim 1, wherein the at least one candidate is generated by using one of maximum a posterior probabilities and maximum likelihood criterion.

18. The method as claimed in Claim 1, wherein the method is a Mandarin directory assistant method.

20 19. A computer program product residing on a computer readable medium having a plurality of instructions stored thereon which, when executed by a processor, cause the processor to:

- receive a speech signal describing the desired directory entry;
- recognize the speech signal and generating recognized word sequences;
- 25 - interpret the recognized word sequences by using a predetermined grammar rule and relevant information thereof stored in a database to form concept sequences;
- interpret the concept sequences according to semantic meaning and relevant information thereof stored in the database and current system status thereof, thereby generating at least one candidate for the desired
- 30 directory entry;

- look up at least one directory entry information corresponding to the at least one candidate from the database; and
- output the at least one directory entry information located.

5 20. A directory assistant apparatus for providing desired directory entry information, comprising:

- a database for storing directory entry information, grammar rules and concept sequences;
- 10 - an acoustic recognition unit for receiving a speech signal describing the desired directory entry, recognizing the speech signal and generating recognized word sequences;
- a speech interpreting unit for interpreting the recognized word sequences by using a predetermined grammar rule and relevant information thereof stored in the database to form concept sequences and interpreting the
- 15 concept sequences according to semantic meaning and relevant information thereof stored in the database and current system status thereof, thereby generating at least one candidate for the desired directory entry;
- a look up unit for looking up at least one directory entry information
- 20 corresponding to the at least one candidate from the database; and
- an output unit for outputting the at least one directory entry information located.

25 21. The directory assistant apparatus of Claim 20, wherein the predetermined grammar rule and concept sequences are used to describe the desired directory entry comprising entry name or at least one word of entry name and relevant information thereof.

30 22. The directory assistant apparatus of Claim 20, wherein the database comprises a relevant knowledge database and a directory database.

23. The directory assistant apparatus of Claim 22, wherein the relevant knowledge database comprises words and using frequencies thereof, ways to describe the words, grammar rules, attributes and corresponding using frequencies.

24. The directory assistant apparatus of Claim 22, wherein the relevant knowledge database comprises communication concepts and frequencies of usage thereof, corresponding grammar rules, semantic meaning and frequencies of usage.

25. The directory assistant apparatus of Claim 22, wherein the directory database comprises a plurality of entries, wherein each entry comprises name, telephone number, relevant information and frequencies of usage.

26. The directory assistant apparatus of Claim 20, wherein the predetermined grammar rule is generated by frequently used grammar templates and frequently used words.

27. The directory assistant apparatus of Claim 26, wherein the grammar templates are generated by one of frequently used nouns, names of famous people, idioms, character strokes, letters, words, and character roots.

28. The directory assistant apparatus of Claim 20, wherein the acoustic recognition unit further comprises a speech recognizer for recognizing the speech signal and generating recognized word sequences.

29. The directory assistant apparatus of Claim 20, wherein the acoustic recognition unit further comprises a confusion analyzer for expanding the recognized word sequences according to a confusion table.

30. The directory assistant apparatus of Claim 29, wherein the confusion table is pre-trained and comprises all confusable words, their corresponding correct ones and occurring probabilities.

31. The directory assistant apparatus of Claim 20, wherein the acoustic recognition unit further comprises a confidence measurement unit for filtering out confusable word pairs according to a confidence table.

32. The directory assistant apparatus of Claim 20, wherein the speech interpreting unit continuously updates the system status.

33. The directory assistant apparatus of Claim 20, further comprising a question generator for generating a question to request more information.

34. The directory assistant apparatus of Claim 33, wherein the question is one of requesting a user to supply more information, listing-based confirmation and open-question confirmation.

35. The directory assistant apparatus of Claim 20, wherein the at least one candidate is generated by using one of maximum a posterior probability and maximum likelihood criterion.

36. The directory assistant apparatus of Claim 20, wherein the outputting unit is a speech output unit.

37. The directory assistant apparatus of Claim 20, wherein the directory assistant apparatus is a Mandarin directory assistant apparatus.

ABSTRACT

DIRECTORY ASSISTANT METHOD AND APPARATUS

5

10

Disclosed is a directory assistant method and apparatus for a large-scale automatic dialogue telecommunications system. The directory assistant method and apparatus use natural language dialogue system to ask users to describe the desired directory entry and then use the relevant knowledge databases to parse and understand these descriptions and interpret their meaning. Finally, the directory assistant method and apparatus integrate all available information from several dialogues turns, directory database and relevant knowledge database, and then provide the users' desired directory entry information.

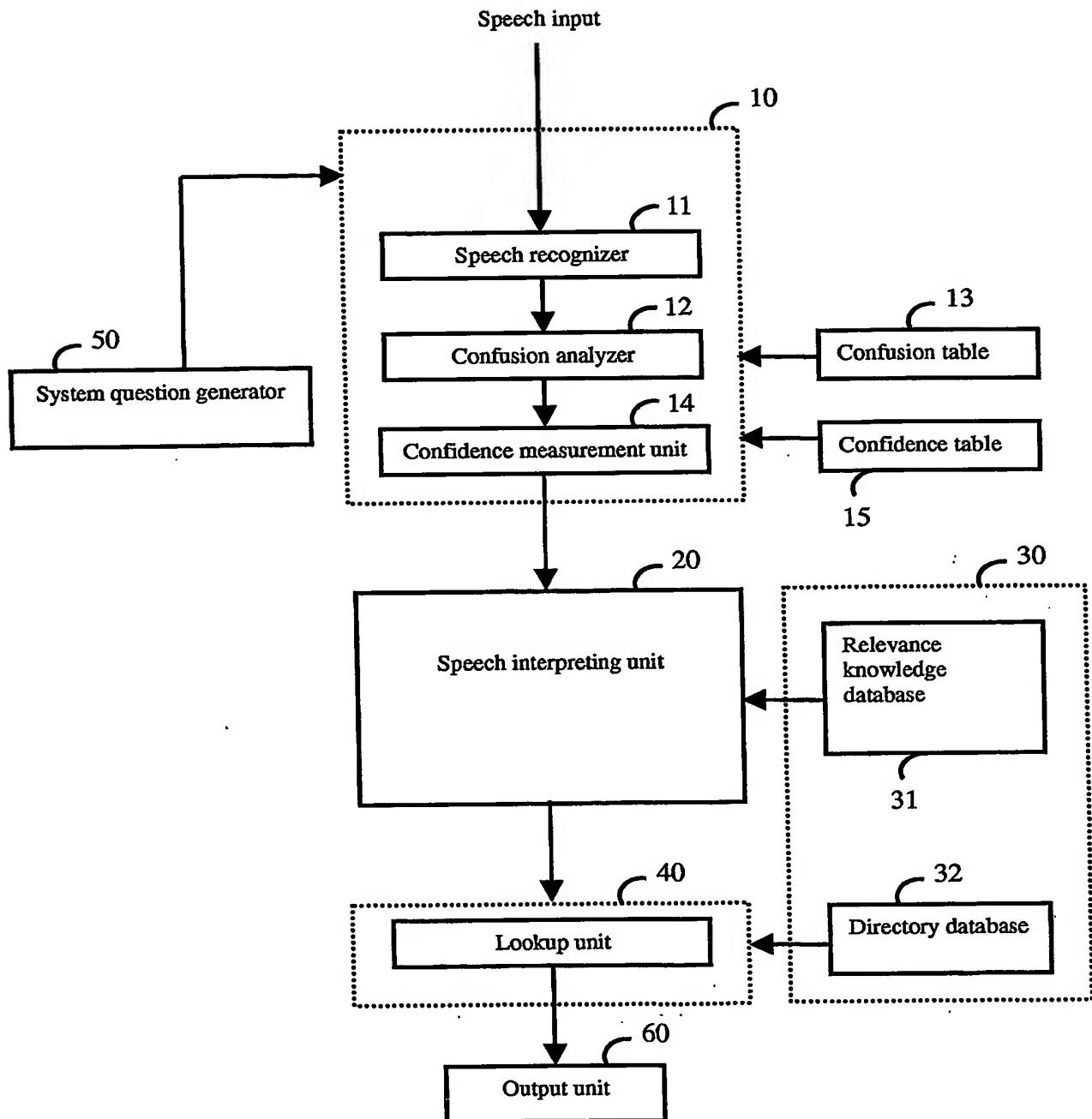


FIG. 1

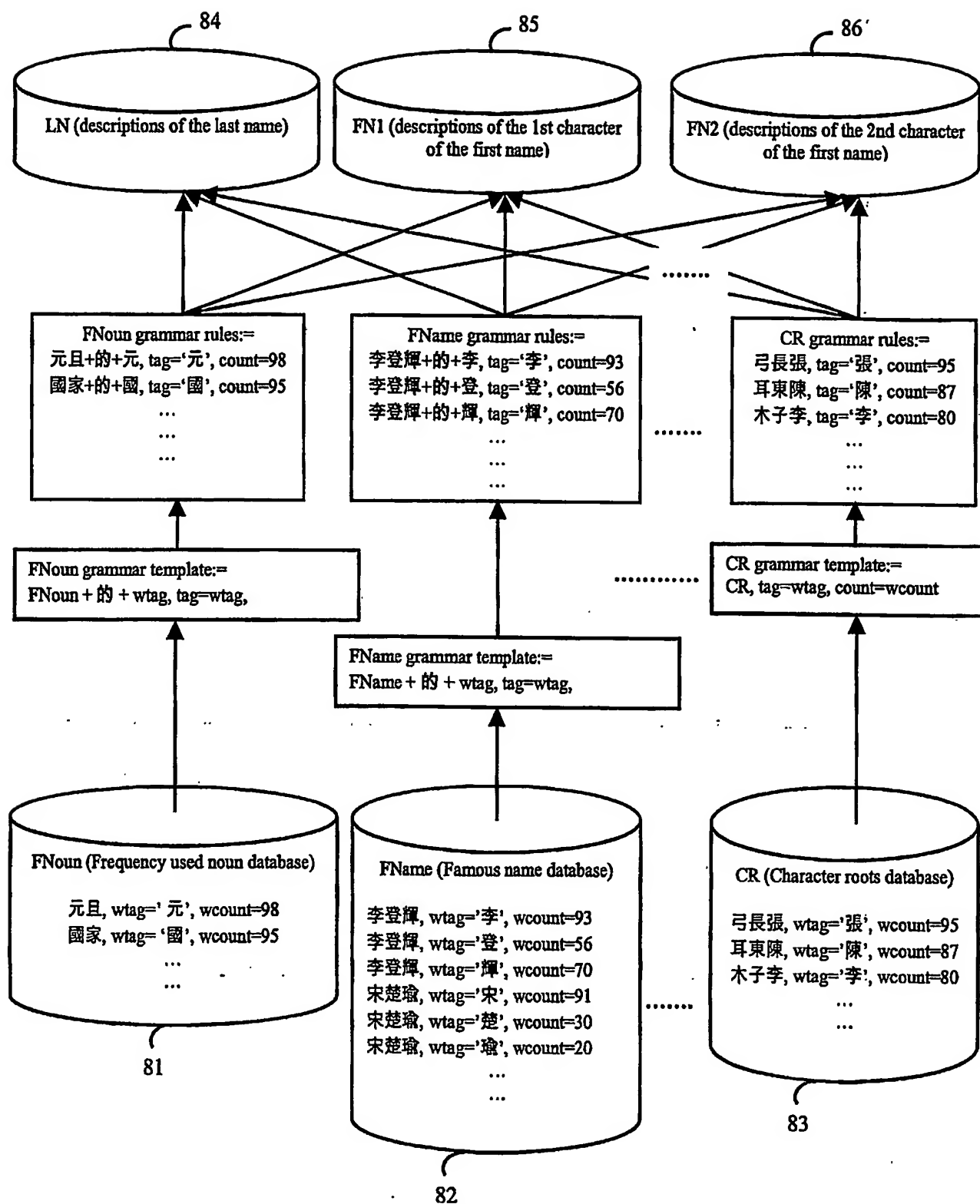


FIG.2

3/3

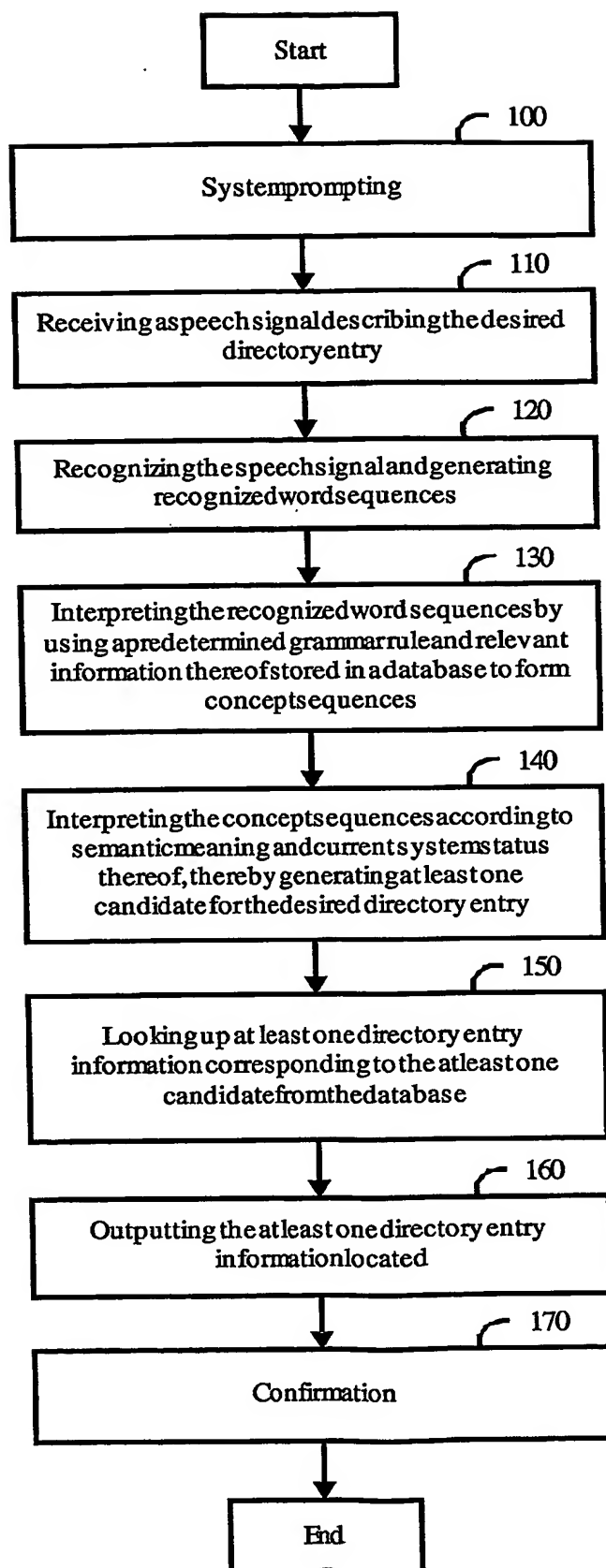


FIG. 3